JOURNAL OF VIROLOGY, July 1996, p. 4707–4713 Vol. 70, No. 7
0022-538X/96/$04.00+0

Characterization of the Rubella Virus Nonstructural Protease Domain and Its Cleavage Site

JUN-PING CHEN,[1] JAMES H. STRAUSS,[2] ELLEN G. STRAUSS,[2] AND TERYL K. FREY[1]*

Department of Biology, Georgia State University, Atlanta, Georgia 30303,[1] and Division of Biology, California Institute of Technology, Pasadena, California 91125[2]

Received 27 October 1995/Accepted 3 April 1996

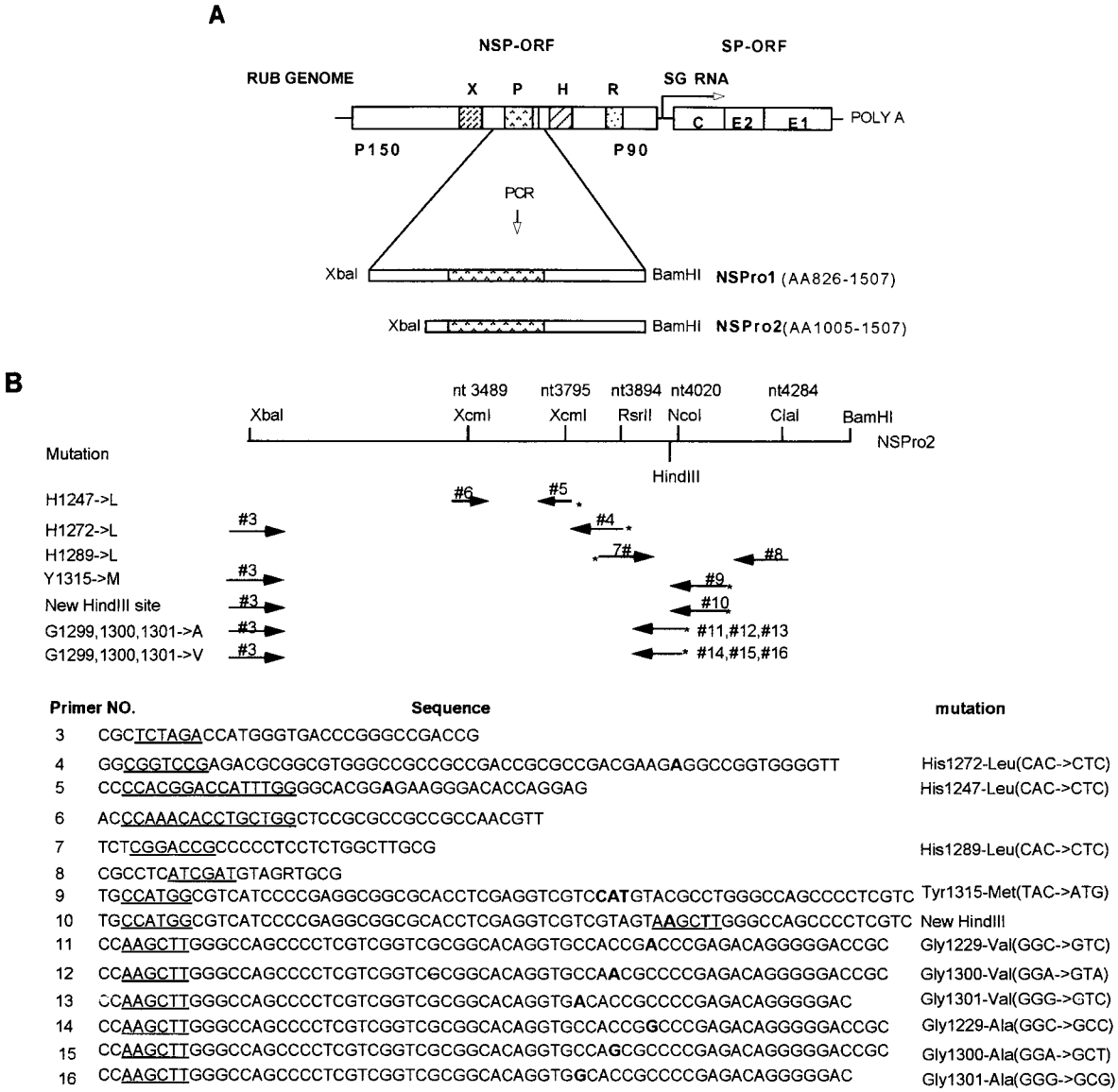
The region of the rubella virus nonstructural open reading frame that contains the papain-like cysteine protease domain and its cleavage site was expressed with a Sindbis virus vector. Cys-1151 has previously been shown to be required for the activity of the protease (L. D. Marr, C.-Y. Wang, and T. K. Frey, Virology 198:586–592, 1994). Here we show that His-1272 is also necessary for protease activity, consistent with the active site of the enzyme being composed of a catalytic dyad consisting of Cys-1151 and His-1272. By means of radiochemical amino acid sequencing, the site in the polyprotein cleaved by the nonstructural protease was found to follow Gly-1300 in the sequence Gly-1299–Gly-1300–Gly-1301. Mutagenesis studies demonstrated that change of Gly-1300 to alanine or valine abrogated cleavage. In contrast, Gly-1299 and Gly-1301 could be changed to alanine with retention of cleavage, but a change to valine abrogated cleavage. Coexpression of a construct that contains a cleavage site mutation (to serve as a protease) together with a construct that contains a protease mutation (to serve as a substrate) failed to reveal *trans* cleavage. Coexpression of wild-type constructs with protease-mutant constructs also failed to reveal *trans* cleavage, even after extended in vitro incubation following lysis. These results indicate that the protease functions only in *cis*, at least under the conditions tested.

Rubella virus (RUB), an enveloped, positive-strand RNA virus, is the sole member of the *Rubivirus* genus of the family *Togaviridae* (for a review, see reference 14). The 9,759-nucleotide genomic RNA contains two long open reading frames (ORFs) organized in a manner similar to the genomes of the members of the genus *Alphavirus*, the other genus in the *Togaviridae* (reviewed in reference 34). The 3′-proximal ORF, 3189 nucleotides in length, encodes the three structural proteins, consisting of a capsid protein (C) and two envelope glycoproteins (E1 and E2). The 5′-proximal ORF, 6615 nucleotides in length, encodes the viral nonstructural proteins (11). Compared with alphaviruses, little is known about the translation and processing of the RUB nonstructural proteins. Analysis of the deduced amino acid sequence of the nonstructural protein ORF (NSP-ORF) revealed the presence of two amino acid motifs conserved among positive-polarity RNA viruses, a GDD motif indicative of replicase activity and a GK(S/T) motif indicative of helicase function. Another RUB motif, of unknown function, X, is present in the genomes of alphaviruses, hepatitis E virus, and coronaviruses but not other RNA viruses. Interestingly, the order of the X motif and the helicase motif in RUB is reversed compared with that in alphaviruses.

Computer-assisted sequence alignment revealed that a region of the RUB NSP-ORF between the X motif and the helicase motif shared significant homology with several viral and cellular papain-like cysteine proteases, including the nsP2 protease which resides within the alphavirus NSP-ORF (16). The catalytic dyad residues of the RUB NS protease predicted from the alignment were Cys-1151 and His-1272. Cleavage of the polypeptide precursor translated from the RUB NSP-ORF was first demonstrated by expression of the NSP-ORF in a vaccinia virus transient-expression system (24). In that study, three NSP-ORF-specific proteins, P200, P150, and P90 (molecular masses of 200, 150, and 90 kDa, respectively), were identified. When one of the predicted catalytic amino acids, Cys-1151, was changed to Gly, P150 and P90 disappeared, demonstrating that a RUB NS protease is responsible for the proteolytic processing of the precursor P200 to P150 and P90 and supporting the hypothesis that Cys-1151 is a component of the catalytic site. Deletion mapping showed that both the protease domain and its cleavage site are located between amino acids 1005 and 1507 of the NSP-ORF. Subsequently, region-specific antisera produced by injection of rabbits with fusion proteins expressed in *Escherichia coli* were used to confirm that in RUB-infected cells only the cleavage of P200 to P150 and P90 could be detected, and the order in the NSP-ORF was found to be NH_2-P150-P90-COOH (13). Thus, the RUB NSP-ORF appears to be cleaved only once by the NS protease, in contrast to alphaviruses, in which the nsP2 protease cleaves three sites in the NSP-ORF precursor.

The goal of this study was to provide support for the hypothesis that His-1272 is a component of the catalytic dyad of the RUB NS protease and to determine the site of cleavage in the NSP-ORF. For many positive-strand RNA viruses, such analyses have been performed with products translated in cell-free systems (8, 12, 32). However, we have been unable to demonstrate proteolytic processing by the RUB NS protease following translation of the NSP-ORF in standard reticulocyte lysates (15). Therefore, for this work, part of RUB NSP-ORF containing the putative protease domain and its cleavage site was expressed in BHK cells by using a Sindbis virus vector (4). Although Sindbis virus vectors have not been used extensively in expression studies, proteins expressed by these vectors are produced in abundance and are processed authentically (7, 27). An advantage of Sindbis virus vectors over the commonly used vaccinia virus vectors is that insertion of the sequences to be expressed is by in vitro ligation rather than by in vivo recombination, simplifying the expression of site-specific mutations made within the expressed sequences. The fragment of the RUB NSP-ORF expressed by the Sindbis virus vector was

* Corresponding author.



Primer NO.	Sequence	mutation
3	CGCTCTAGACCATGGGTGACCCGGGCCGACCG	
4	GGCGGTCCGAGACGCGGCGTGGGCCGCCGCCGACCGCGCCGACGAAGAGGCCGGTGGGGTT	His1272-Leu(CAC->CTC)
5	CCCCACGGACCATTTGGGGCACGGAGAAGGGACACCAGGAG	His1247-Leu(CAC->CTC)
6	ACCCAAACACCTGCTGGCTCCGCGCCGCCGCCAACGTT	
7	TCTCGGACCGCCCCCTCCTCTGGCTTGCG	His1289-Leu(CAC->CTC)
8	CGCCTCATCGATGTAGRTGCG	
9	TGCCATGGCGTCATCCCCGAGGCGGCGCACCTCGAGGTCGTCCATGTACGCCTGGGCCAGCCCCTCGTC	Tyr1315-Met(TAC->ATG)
10	TGCCATGGCGTCATCCCCGAGGCGGCGCACCTCGAGGTCGTCGTAGTAAGCTTGGGCCAGCCCCTCGTC	New HindIII
11	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGCCACCGACCCGAGACAGGGGGACCGC	Gly1229-Val(GGC->GTC)
12	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGCCAACGCCCCGAGACAGGGGGACCGC	Gly1300-Val(GGA->GTA)
13	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGACACCGCCCCGAGACAGGGGGAC	Gly1301-Val(GGG->GTC)
14	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGCCACCGGCCCGAGACAGGGGGACCGC	Gly1229-Ala(GGC->GCC)
15	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGCCAGCGCCCCGAGACAGGGGGACCGC	Gly1300-Ala(GGA->GCT)
16	CCAAGCTTGGGCCAGCCCCTCGTCGGTCGCGGCACAGGTGGCACCGCCCCGAGACAGGGGGAC	Gly1301-Ala(GGG->GCG)

FIG. 1. Regions of RUB NSP-ORF expressed by dsSIN recombinants and mutagenesis strategy. (A) In the diagram of the RUB genome, ORFs and nontranslated regions are represented by open boxes and lines, respectively (NSP-ORF, nonstructural protein ORF; SP-ORF, structural protein ORF). Within the NSP-ORF, the X, protease, helicase, and replicase domains are indicated as X, P, H, and R, respectively. The boundaries of the NSPro1 and NSPro2 fragments, which were generated by PCR and introduced into the dsSIN vector, are shown. AA, amino acids. (B) Oligonucleotide-directed, site-specific mutagenesis was done by PCR by incorporating the mutation into one member of the primer pair. The amplification product containing the mutation was introduced into pH2J1-NSPro2 at convenient restriction sites. At the top is a map of the NSPro sequence showing the primer pairs used in each mutagenesis and restriction sites used to introduce mutations into pH2J1-NSPro2. The arrows denote the orientation of each primer; the stars indicate the primer that contained the mutation. One mutation created a *Hin*dIII site which was used in subsequent mutagenesis. At the bottom, the sequences of the primers are given; underlines denote restriction sites, and letters in boldface type denote the changed nucleotides (nt).

found to self-cleave, allowing site-specific mutagenesis and sequence analysis to be performed.

MATERIALS AND METHODS

Plasmid construction. The dsSIN expression vector, Toto2JC1, and the shuttle vector, pH2J1 (4), were provided by Charles Rice. Since the RUB genome contains several *Xho*I sites, a unique *Mlu*I site was introduced in both plasmids immediately downstream of the *Xho*I site normally used for linearization for runoff transcription. Two fragments from the RUB genome, NSPro1 (nucleotides 2515 to 4560) and NSPro2 (nucleotides 3053 to 4560), were amplified by PCR from Robo12, a plasmid containing a cDNA copy of the RUB genome (35) (shown schematically in Fig. 1A). For production of the NSPro1 fragment, the upstream primer (primer 1) was CGCTCTAGATC**ATG**ATGGACCCACCGCC CGGC and the downstream primer (primer 2) was CGCGGATCC**TACTA**GAC CTGGCGGCCGTCCCA. For production of the NSPro2 fragment, the upstream primer (primer 3) was CGCTCTAGACC**ATG**GGTGACCCGGGCC GACCG and the downstream primer was primer 2. The primers contained restriction sites for cloning, *Xba*I in primers 1 and 3 and *Bam*HI in primer 2 (underlined in the primer sequences), as well as new start codons in the upstream primers and new stop codons in the downstream primer (shown in boldface type in the primer sequences) to allow translation of appropriate products from these internal fragments. The NSPro1* and NSPro2* fragments were generated with the same primer sets and with pTM3/nsRUB* as a template; pTM3/nsRUB* is a plasmid that contains a copy of the RUB NSP-ORF in which Cys-1151 is changed to Gly (24). The NSPro1, NSPro1*, NSPro2, and NSPro2* PCR products were digested with *Xba*I and *Bam*HI and introduced into pH2J1 between the *Xba*I and *Bam*HI sites in the polylinker. Subsequently, the *Xba*I-*Mlu*I fragments from the pH2J1-NsPro constructs were transferred to Toto2JC1 between the *Xba*I and *Mlu*I sites. All PCRs were performed in a volume of 200 μl in the

manufacturer's (Amersham Life Science) buffer (50 mM Tris-HCl [pH 9.0], 1.5 mM $MgCl_2$, 0.02 mM ammonium sulfate), containing 200 μM each deoxynucleoside triphosphate (dNTP), 1 μM each primer, 20 μl of dimethyl sulfoxide, 16 μl of bovine serum albumin, 7 μl of 1% 2-mercaptoethanol, and 5 U of Hot tub DNA polymerase (Amersham Life Science), for 30 cycles (94°C for 1 min, reduction to 55°C over 2 min, increase to 72°C over 2 min, 72°C for 3 min) followed by a 10-min extension at 72°C in a Perkin-Elmer Cetus thermal cycler. For recombinant plasmid constructions, standard recombinant DNA techniques as described by Sambrook et al. (28) were used with minor modifications.

Site-directed mutagenesis of the RUB NSP sequence. Oligonucleotide-directed site-specific mutagenesis was done by PCR by incorporating the mutation into one of the pair of primers. The product amplified with these primers from a Robo12 template was introduced into pH2J1-NSPro2 by using convenient restriction sites followed by transfer into Toto2JC1. The PCR fragment was sequenced after introduction into either pH2J1 or Toto2JC1 to confirm the presence of the mutation introduced and to ensure that no other mutations were introduced during PCR amplification. A diagram of the strategy and the sequences of the oligonucleotides used in mutagenesis are shown in Fig. 1B. No suitable restriction site was present near the nucleotides encoding the cleavage site, and for mutagenesis of the cleavage site a unique *Hin*dIII site (AAGCTT) was created near the cleavage site by introducing two silent changes (CAGGCGT to CAAGCTT).

Production of recombinant dsSIN viruses. BHK-21 cells were grown at 35°C under 5% CO_2 in Dulbecco's modified Eagle's medium (Gibco/BRL) containing 10% fetal bovine serum and 50 μg of gentamicin per ml. To produce stocks of recombinant viruses, 3 to 4 μg of each Toto2JC1 recombinant plasmid was linearized with *Mlu*I and transcribed in vitro with SP6 polymerase in the presence of cap analog (35). An aliquot of the transcription reaction mixture was used to transfect 5×10^6 BHK cells by electroporation (23). The supernatant was harvested at 24 h posttransfection. Titers of recombinant virus stocks ranged from 5×10^8 to 1×10^9 PFU/ml. For dsSIN recombinants containing mutagenized NSPro inserts, the presence of the mutation in the virus was confirmed as follows. Recombinant virus stock (250 to 500 μl) was extracted with Tri-Reagent LS (Molecular Research Center) as specified by the manufacturer, and virion RNA was collected by isopropanol precipitation. The virion RNA was used in a reverse transcription-PCR with a 100-μl volume of mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5 mM $MgCl_2$, 10 mM dithiothreitol, 1.25 mM each dNTP, 1 μM each primer (the upstream primer was primer 3; the downstream primer was CTGATATTGTAAGGGCC [complementary to nucleotides 4064-4080 of the RUB genome]), 2 U of RNasin RNase inhibitor (Promega), 20 U of avian myeloblastosis virus reverse transcriptase (Molecular Genetic Resources), and 2.5 U of *Taq* polymerase (Perkin Elmer-Cetus). The reverse transcription-PCR mixture was incubated at 42°C for 60 min and 95°C for 3 min and then subjected to 35 cycles of 95°C for 1 min, 55°C for 2 min, and 72°C for 3 min, followed by an extension at 72°C for 10 min. Alternatively, the reverse transcription reaction was completed before addition of the *Taq* polymerase. The PCR product, which encompasses nucleotides 3053 to 4080 of the RUB genome, was gel purified before being sequenced.

Protein preparation and analysis. Confluent monolayers of BHK cells in six-well culture plates (containing about 10^7 cells per well) were infected with each recombinant dsSIN virus at a multiplicity of infection of 10 PFU per cell. At 16 h postinfection, the medium was removed and replaced with Dulbecco's modified Eagle's medium lacking methionine and containing 2% dialyzed fetal bovine serum. Thirty minutes later, ^{35}S-Translabel (1,108 Ci/mmol; ICN) or [^{35}S]methionine (1,150 Ci/mmol; Amersham) was added to the medium at 50 μCi/ml. The cells were labeled continuously for 1 h or labeled for 30 min and then chased with Dulbecco's modified Eagle's medium containing 10% fetal bovine serum for 1, 2, or 3 h. Following the radiolabeling protocol, the medium was removed and the cells were rinsed three times with phosphate-buffered saline and solubilized in 0.5 ml of RIPA buffer (1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 150 mM NaCl, 3 mM EDTA, 10 mM Tris [pH 7.6]) containing 50 μg of antipain-dihydrochloride (Boehringer Mannheim) per ml or TNE buffer (10 mM Tris-HCl [pH 7.6], 150 mM NaCl, 3 mM EDTA) containing 1% Nonidet P-40. Following removal of nuclei by centrifugation, the supernatants were mixed 1:1 with 2× Laemmli sample buffer (50 mM Tris-HCl [pH 6.8], 3% SDS, 10% glycerol, 1% 2-mercaptoethanol, 0.025% bromphenol blue), boiled for 3 min, and electrophoresed in SDS-12% polyacrylamide gels (22). Alternatively, lysates were subjected to immunoprecipitation followed by SDS-polyacrylamide gel electrophoresis (PAGE) (13).

Amino acid sequencing. BHK cells (5×10^7) in 60-mm culture plates were infected with dsSIN/NSPro2 or dsSIN/NSPro2:Y1315M (a recombinant in which the tyrosine residue at 1315 was replaced with methionine) at a multiplicity of infection of 10 PFU per cell. At 16 h postinfection, the cells were radiolabeled for 3 h in medium containing 100 μCi of [^{35}S]methionine (1,150 Ci/mmol; Amersham Corp.) per ml or 100 μCi of [^{35}S]methionine per ml and 50 μCi of [^{3}H]leucine per ml (47 Ci/mmol; Amersham Corp.). Following immunoprecipitation, the protein A-Sepharose-immune complex conjugates were dissociated in 1% SDS–5% 2-mercaptoethanol–62.5 mM Tris (pH 7.6) at 100°C for 3 min, and the supernatant was concentrated by rotary evaporation and resolved by SDS-PAGE in a 15% polyacrylamide gel. The contents of the gel were electroblotted onto a polyvinylidene difluoride membrane (Applied Biosystems) with a Bio-Rad apparatus at 60 V at room temperature for 1.5 h in buffer containing 25 mM Tris base (pH 8.3), 192 mM glycine, and 10% (vol/vol) methanol. The 24-kDa C-terminal cleavage product was localized by autoradiography and excised from the polyvinylidene difluoride blot. Radioactive sequence analysis was carried out by G. Hathaway with a model 477A protein sequencer (Perkin Elmer/Applied Biosystems Division). Samples were sequenced directly from the polyvinylidene difluoride membrane by a modification of the normal Edman cycle together with the instrument's microblot cartridge. The radioactive amino acids were assayed as their anilinothiazolinone derivatives by redirecting the effluent from the reaction cartridge to a fraction collector, and fractions corresponding to each cycle were assayed for radioactivity in a liquid scintillation counter.

RESULTS

Construction of dsSIN viruses that express the RUB NS protease domain. The RUB NS protease domain and its cleavage site have been shown by deletion mapping to be located within the domain composed of amino acids 1005 to 1507 (24). Two regions of the RUB NSP-ORF containing this region (NSPro1 [amino acids 826 to 1507] and NSPro2 [amino acids 1005 to 1507]) were introduced into the SIN expression vector Toto2JC1; in this double-subgenomic (ds) vector, the RUB sequences are expressed from a distinct mRNA transcribed from the second subgenomic promoter (reviewed in reference 4). In dsSIN/NSPro1- or dsSIN/NSPro2-infected cell lysates, three new proteins (Fig. 2A) were observed in addition to those in SIN-infected cell lysates, and these proteins were immunoprecipitated by specific anti-RUB NSP-ORF serum (Fig. 2B). The NSPro1 proteins had apparent molecular masses of 76, 52, and 24 kDa, whereas the NSPro2 proteins has apparent masses of 56, 32, and 24 kDa. Two mutant constructs, dsSIN/NSPro1* and dsSIN/NSPro2*, in which the catalytic Cys-1151 was changed to Gly, synthesized only the 76- and 56-kDa proteins, respectively. These species therefore are the full-length translation products, and the smaller NSP-ORF-specific proteins are cleavage products. Since the NSPro2 sequence was truncated at its N terminus by 179 amino acids in relation to NSPro1, leading to the reduction in the size of the precursor from 76 kDa in NSPro1 to 56 kDa in NSPro2, a similar size reduction should occur in the N-terminal cleavage product of NSPro2 in comparison with NSPro1 whereas the sizes of the C-terminal cleavage product of both constructs should be similar. Thus, the 24-kDa cleavage product must be the C-terminal cleavage product whereas the 52-kDa product from NSPro1 and the 32-kDa product from NSPro2 are the N-terminal products. In addition, in both dsSIN/NSPro2- and dsSIN/NSPro2*-infected cells, a protein of 52 to 53 kDa was immunoprecipitated by NSP-ORF-specific antisera (Fig. 2B). The identity of this protein is unknown; however, because it was produced by dsSIN/NSPro2*, its presence was not due to activity of the RUB NS protease.

In the experiment in Fig. 2, dsSIN/NSPro-infected cells were labeled for 1 h. In pulse-chase experiments, maximal product/precursor ratios were reached after a 1-h chase and maintained during a 3-h chase but significant amounts of precursor remained (data not shown). Thus, cleavage occurred only directly after translation and was not complete.

Mutagenic analysis of the RUB protease domain. In general, papain-like proteases have a catalytic dyad consisting of one cysteine and one histidine residue. On the basis of computer alignment, Cys-1151 and His-1272 in the RUB NSP-ORF were proposed to form the catalytic dyad (16). We previously tested the importance of Cys-1151 (24). We have now tested the importance of His-1272, as well as that of the two neighboring His residues, His-1247 and His-1289, by changing each individually to leucine. As shown in Fig. 3, mutation of His-1272 completely abolished proteolytic activity whereas mutation of either His-1247 or His-1289 had no effect. These results sup-

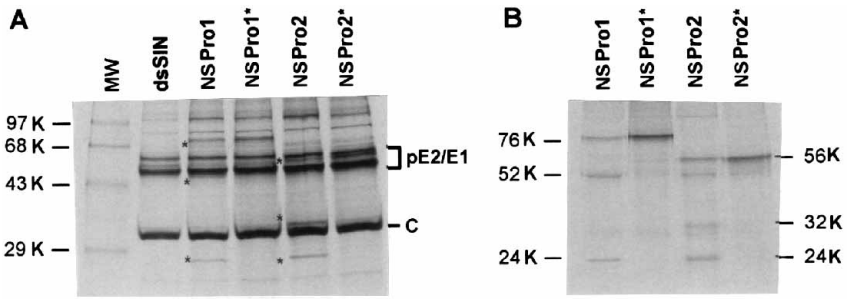


FIG. 2. Expression of NSPro sequences by dsSIN viruses. BHK cells infected with the dsSIN vector or dsSIN recombinants (multiplicity of infection, 10 PFU per cell) were labeled for 1 h with [^{35}S]methionine at 16 h postinfection, after which the cells were solubilized with 1× RIPA buffer. (A) Infected-cell lysates were resolved by SDS-PAGE. The left-hand lane contains molecular weight (MW) standards (denoted in the left margin in thousands). The SIN structural proteins are denoted in the right margin. The NSPro products are indicated by stars. (B) SDS-PAGE following immunoprecipitation of the infected cell lysates with GU-6, an antiserum against a fusion protein containing the region of the NSP-ORF present in NSPro (13). The apparent molecular weights (in thousands) of the immunoprecipitated products are indicated. Lanes: NSPro1 and NSPro2, lysates of cells infected with dsSIN/NSPro1 and dsSIN/NSPro2, respectively; NSPro1* and Pro2*, lysates of cells infected with mutant constructs in which Cys-1151 was replaced with Gly. Autoradiographs of PAGE gels were scanned with an Agfa Arcus II flatbed scanner operating with Fotolook V2.05 under Adobe Photoshop V2.5 software. Prints were made with a Tektronics Phaser 440 dye sublimation printer with a monochrome ribbon.

port the hypothesis that Cys-1151 and His-1272 form a catalytic dyad in the protease.

Determination of the cleavage site. Radiochemical sequencing was used to identify the cleavage site in the RUB NS polyprotein. From the electrophoretic mobilities of the dsSIN/NSPro2 cleavage products, the cleavage site should be between residues 1290 and 1310, just downstream of the catalytic His-1272. The cleavage sites of most viral papain-like proteases follow amino acids with small side chains, such as glycine and alanine (12, 17), and this region contained several candidate GG and AA dipeptides. In one experiment, cells infected with a construct in which Tyr-1315 was changed to methionine (dsSIN/NSPro2:Y1315M) (this substitution had no effect on the activity of the protease) were labeled with [^{35}S]methionine and the 24-kDa cleavage product was isolated and sequenced. Radiolabeled peaks were observed in cycles 15 and 28 (Fig. 4A), which could be aligned with methionines at residues 1315 and 1328, indicating that cleavage occurs after Gly-1300. To confirm the cleavage site, dsSIN/NSPro2 recombinant virus-infected cells were doubly labeled with [^{35}S]methionine and [^{3}H]leucine. Edman degradation of the 24-kDa cleavage product yielded tritium peaks in cycles 10, 18, and 23 and a ^{35}S peak in cycle 28 (Fig. 4B), which aligned with leucines at residues 1310, 1318, and 1323, as well as with methionine at residue 1328 (Fig. 4B). Note that the ^{35}S peak in cycle 15 in Fig. 4A was not present, since dsSIN/NSPro2 expresses the 24-kDa protein with tyrosine at residue 1315 rather than methionine. These results unambiguously identify Gly-1301 as the N terminus of the 24-kDa product.

Characterization of the RUB NS protease cleavage site. The amino acid sequence P2-P1-P1′ surrounding the RUB cleavage site is Gly-1299–Gly-1300–Gly-1301. It is noteworthy that the corresponding residues in alphaviruses are Gly-Gly/Ala/Cys-X, where X can be any of several amino acids (34). To test the requirements for the three glycine residues in RUB, they were each changed individually to alanine or to valine. As shown in Fig. 5, changing any of the three glycines to valine resulted in the complete abolition of cleavage (lanes V2, V1, and V1′). In contrast, changing glycine to alanine at the P2 and P1′ positions impaired but did not eliminate cleavage (lanes A2 and A1′) whereas this substitution at the P1 position resulted in only minimally detectable cleavage (lane A1). Thus, the RUB cleavage site is most sensitive to the identity of the P1 residue, but the P2 and P1′ residues also contribute significantly to cleavage site recognition.

Assay for *trans* activity of the RUB NS protease. To examine whether the RUB NS protease can act in *trans*, the dsSIN/NSPro2* mutant (which expresses a nonfunctional protease) and the dsSIN/NSPro2:G1300V mutant (which expresses a noncleavable cleavage site) were used to coinfect BHK cells at a multiplicity of infection of 20 PFU per cell for each recombinant. Indirect-immunofluorescence experiments demonstrated that when cells were infected individually with each recombinant virus at this multiplicity, almost 100% of cells expressed the RUB protein products (data not shown). In coinfected cells, the functional protease expressed by dsSIN/NSPro2:G1300V should act on the unimpaired cleavage site of the dsSIN/NSPro2* product if the protease can act in *trans*. However, as shown in Fig. 6A, only the 56-kDa precursor protein was observed in coinfected cells and no cleavage products were detectable (lane NSPro2*+G1300V), indicating that under these conditions, the protease does not function in *trans.*

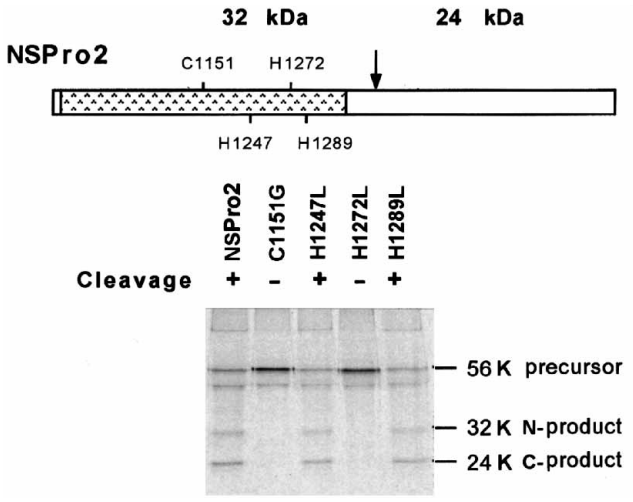


FIG. 3. Identification of the catalytic His residue in the RUB NS protease. Single amino acid substitutions were introduced into NSPro2 at the positions shown in the schematic diagram at the top of the figure, and dsSIN/NSPro2 recombinants containing the mutations were generated. BHK cells infected with dsSIN/NSPro2 or NSPro2 mutants were radiolabeled, lysed, and immunoprecipitated with GU-6 prior to resolution by SDS-PAGE, as shown at the bottom of the figure. In the line marked cleavage, + indicates that cleavage occurred and − indicates that the cleavage was abolished.

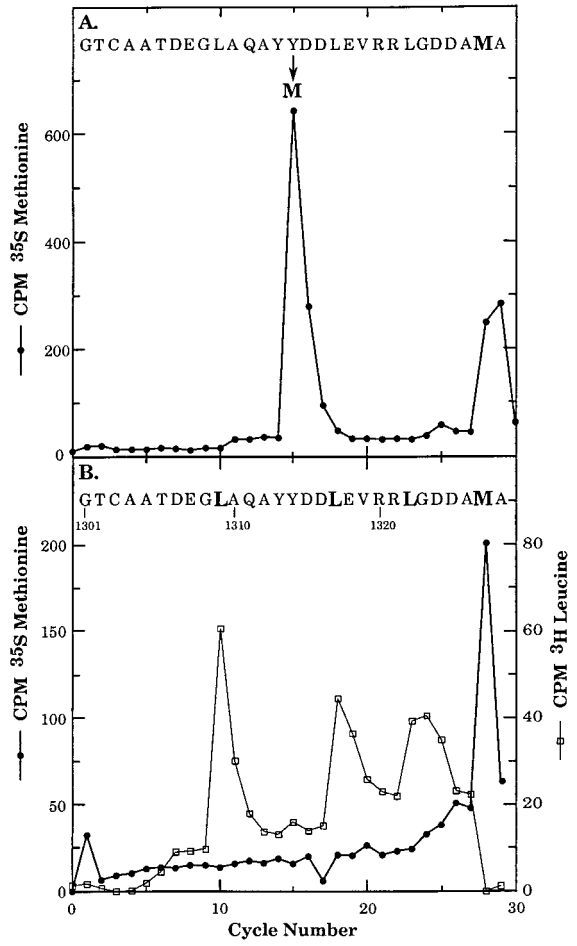


FIG. 4. Sequencing of the NSPro2 C-terminal cleavage product. The 24-kDa cleavage product from dsSIN/NSPro2:Y1315M-infected cells labeled with [^{35}S]methionine (A) or dsSIN/NSPro2-infected cells doubly labeled with [^{35}S]methionine and [^{3H}]leucine (B) was isolated and subjected to Edman degradation, and the radioactivity released during each cycle was determined. The amino acid sequences of NSPro2:Y1315M (A) and NSPro2 (B) are shown at the top of the panels.

As an additional approach to detect *trans* cleavage, BHK cells were coinfected with dsSIN/NSPro1 and dsSIN/NSPro2* or with dsSIN/NSPro1* and dsSIN/NSPro2. Both of these coinfections were with one wild-type construct and one protease mutant construct. Since the sizes of the N-terminal cleavage product of the NSPro1 and NSPro2 constructs differ (Fig. 2), *trans* cleavage of the protease mutant template by the wild-type construct would be detectable. However, as shown in Fig. 6B (lanes Pro1+Pro2* and Pro1*+Pro2), in both coinfections only the N-terminal product from *cis* cleavage by the wild-type construct was present. Lysates from coinfected cells (in either RIPA buffer or TNE-1% Nonidet P-40) were incubated for up to 24 h at both room temperature and 35°C, with the result that no change in the cleavage pattern could be observed. Thus, as in the previous experiment, we were unable to detect *trans* cleavage.

DISCUSSION

We have shown that Cys-1151 and His-1272 are required for catalytic activity by the RUB NS protease, consistent with the prediction of Gorbalenya et al. (16) that these residues form the catalytic dyad of a papain-like protease. Additional evidence that the RUB NS protease belongs to the papain family is that it is inhibited by E-64d, a specific inhibitor of thiol protease activity (6a, 18, 20, 25). Other viral proteases belonging to the viral papain-like protease family are the alphavirus nsP2 protease, the equine arteritis virus (EAV; family *Arteriviridae*) nsP1 protease, the murine hepatitis virus (MHV; family *Coronaviridae*) PLP-1, the potyvirus HC-Pro, the p29 and p48 proteases of hypovirulence-associated virus (HAV) of chestnut blight fungus, and possibly the leader protease of foot-and-mouth disease virus (FMDV; genus *Aphthovirus*, family *Picornaviridae*) (1, 2, 8, 12, 26, 29, 31, 32).

We also found that the cleavage site of the RUB NS protease followed Gly-1300, 28 residues downstream from the catalytic His residue. An alignment of cleavage sites recognized by viral papain-like proteases is shown in Fig. 7. The most uniform feature of these sites is the presence of an amino acid with a short side chain in the P1 position. The P1′ position is also usually occupied by an amino acid with a short side chain. In particular, Gly-Gly is present in the P1-P1′ position not only in the RUB site but also in the EAV nsp1, potyvirus HC-Pro, and HAV p29 sites. The RUB and alphavirus NS cleavage sites share a Gly in the P2 position, whereas the sites of the other viral proteases have bulky hydrophobic amino acids or arginine in the P2 position. Notably, the Gly in the P2 position is the only residue that is strictly conserved in all 10 alphaviruses whose genomes have been sequenced (34). As is the case for Gly in the P2 position of the RUB site, substitution of the P2 Gly by alanine in the alphavirus cleavage sites is tolerated, although the cleavage efficiency is reduced, but substitution by other amino acids results in a site completely refractory to cleavage. Studies of the cleavage sites of the potyvirus HC-Pro, the HAV p29 and p48 proteases, and the MHV PLP-1 have also shown that the P2 residue is important for substrate recognition by these proteases (5, 6, 12, 17, 29). Although the P2 and P1 residues are similar between RUB and alphavirus sites, the nature of the residues at the P3 and P1′ positions is different. An Ala is usually present at the alphavirus P3, but Arg is present in the RUB site. Alphaviruses tolerate amino acids with bulky side chains in the P1′ position (a Tyr in naturally present in the nsP3/nsP4 cleavage site of all alphaviruses, and Phe, Met, and Arg are also tolerated in this position [9, 10, 30]), whereas the RUB site is quite sensitive to any change in the glycine present at P1′. Similar to the RUB cleavage site, the cleavage sites for the EAV nsP1 and potyvirus HC-Pro proteases cannot tolerate substitution of the P1′ Gly with a bulky amino acid (6, 31).

We were not able to detect *trans* cleavage by the RUB NS

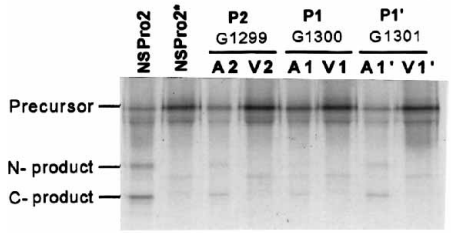


FIG. 5. Mutagenesis of amino acids around the cleavage site. dsSIN/NSPro2 recombinants in which Gly in the P2, P1, or P1′ position of the cleavage site (amino acids 1299, 1300, and 1301) was substituted individually with Ala (A2, A1, or A1′, depending on whether the substitution was made at the P2, P1 or P1′ position) or Val (V2, V1, or V1′) were constructed. BHK cells infected with dsSIN/NSPro2 (NSPro2), dsSIN/NSPro2* (NSPro2*) or the cleavage site mutants were radiolabeled, lysed, and immunoprecipitated with GU-6, prior to resolution by SDS-PAGE. The 56-kDa precursor and the 32- and 24-kDa N- and C-terminal cleavage products are indicated.

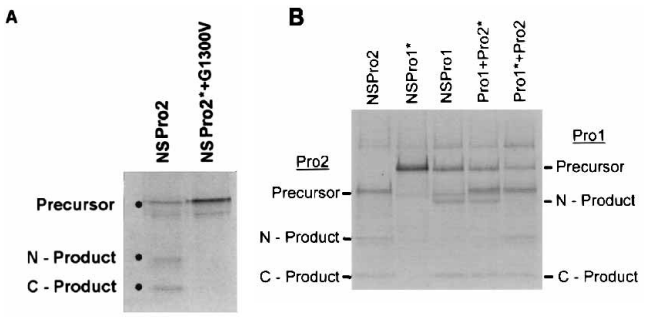


FIG. 6. Assay for cleavage in *trans*. (A) BHK cells were infected with dsSIN/NSPro2 (lane NSPro2) or coinfected with dsSIN/NSPro2* (mutated catalytic site) and dsSIN/NSPro:G1300V(V1) (mutated cleavage site) (lane NSPro2*+G1300V). (B) BHK cells were infected with dsSIN/NSPro2 (lane NSPro2), dsSIN/NSPro1 (lane NSPro1), or dsSIN NSPro1* (lane NSPro1*) or coinfected with dsSIN/NSPro1 and dsSIN/NSPro2* (lane Pro1+Pro2*) or with dsSIN/NSPro1* and dsSIN/NSPro2 (lane Pro1*+Pro2). In both experiments, after radiolabeling had been carried out for 1 h at 16 h postinfection, cell lysates were prepared, immunoprecipitated with GU-6, and resolved by SDS-PAGE. In both panels, the precursor and cleavage products are denoted.

protease. It remains to be ascertained if *trans* activity is also lacking when the protease is produced in its natural context of the P200 precursor or P150 product. If, as it appears, the RUB NS protease lacks *trans* activity, however, then it differs from the activity of the alphavirus nsP2 protease. Interestingly, members of the viral papain-like protease family have been differentiated into two classes on the basis of their ability to function in *cis* or *trans* (16). Main or M-proteases exhibit *trans* activity and cleave at several sites within the viral polyprotein precursor, whereas leader or L-proteases cleave in *cis* at a single site. Thus far, the alphavirus nsP2 protease is the only defined M-protease within the papain-like protease family, although a number of viral proteases which are not papain-like proteases exhibit this general type of activity (e.g., the picornavirus $3C^{pro}$ cysteine protease, the flavivirus NS3 serine protease, and the arterivirus serine protease [33]). The FMDV leader protease also appears to be an M-protease, since it cleaves its single site within the viral polyprotein precursor in *trans* and also cleaves sites within cellular proteins in *trans* (19, 26). L-proteases, which include the EAV nsP1, the potyvirus HC-Pro, and the HAV p29 proteases, cleave in *cis* at a single site, which in each case is a Gly-Gly dipeptide situated 31 to 42 amino acids from the catalytic His residue. In contrast, the single site cleaved in *cis* by the MHV PLP-1 is upstream from the protease domain, and thus this enzyme differs from the other L-proteases (17). There is evidence that this protease cleaves at a second site in the MHV ORF-1 precursor and thus may be an M-protease (3).

By these criteria, the RUB NS protease appears to belong to the L-protease category, although Gorbalenya et al. (16) predicted that it would be an M-protease. However, the activity of the RUB NS protease differs from that of the other L-proteases in three regards. First, the RUB NS protease domain is in the center of a long ORF and mediates the cleavage of a polyprotein precursor into two mature products. In contrast, the other L-proteases are located at the N termini of the ORFs in which they reside and function to cleave off relatively short "leader proteins." The ORFs encoding the EAV nsP1 protease, the potyvirus HC-Pro, and the MHV PLP-1 encode other proteases, which mediate more extensive processing of the translation product of the ORF. Second, while the other L-proteases cleave cotranslationally, the RUB NS protease cleavage occurs after translation of the NSP-ORF is complete (13). Third, the cleavage activity of the other L-proteases occurs efficiently in vitro in a standard rabbit reticulocyte translation reaction, whereas the in vitro activity of the RUB NS protease is detectable only in the presence of a HeLa cell extract (15).

Virus	Protease	Site	P4	P3	P2	P1	↓	P1'	P2'	P3'	P4'
MHV	PLP-1		G	Y	R	G		V	K	P	I
EAV	nsP1		G	N	Y	G		G	Y	N	P
Potyvirus	HC-Pro		Y	x	V	G		G	x	x	x
HAV	p29		A	R	I	G		G	R	L	N
	p48		I	L	V	G		A	E	E	G
RUB	NSPro		S	R	G	G		G	T	C	A
Alphavirus	nsP2	nsP1/2	x	A	G	A		A/G	x	V	E
		nsP2/3	x	A	G	A/C/R		A	P	S	Y
		nsP3/4	x	V/A	G	A/G		Y	↓	E	S
FMDV	L-Pro	L/VP4	R	K	L	K		G	A	G	N
		eIF	A	N	L	G		R	P	A	L

FIG. 7. Comparison of the amino acid sequence around cleavage sites of viral papain-like proteases. The P4 to P4′ residues of the cleavage sites of the recognized papain-like proteases of positive-strand RNA viruses are shown. Sequences are from MHV PLP-1 (12, 17), EAV nsP1 (31), HAV p28 and p48 (8, 29), and FMDV (strain A12; L-VP4 cleavage is within the viral polyprotein precursor; the eIF site is in the cellular translation factor eIF-4γ) (19, 26). The potyvirus HC-Pro cleavage site is a consensus of sequences from around 20 potyviruses (4, 4a), and the alphavirus nsP2 cleavage sites are the consensus from 10 alphavirus sequences (34). In the consensus sequences, underlined amino acids are invariant and X denotes no conservation at that position.

Although RUB and the alphaviruses have a similar genomic expression strategy and are classified in the same family, the order of conserved motifs is different within the NSP-ORFs (11), and phylogenetic analyses show that the helicase and replicase domains of RUB and the alphaviruses are more closely related to those of other viruses than to each other (21, 36). The results of this study suggest that the NS proteases of these viruses differ in functional activity (which would explain the difference in NSP processing exhibited by these viruses). It is possible that these differences result from a different origin of the RUB and alphavirus NS regions. It would be of interest to determine if the RUB NS protease was more similar to viral L-proteases, whose activities are similar to the RUB NS protease, or to the alphavirus NS protease. Unfortunately, the viral papain-like proteases are too widely divergent to permit a phylogenetic analysis of their linear sequences to answer this question (16). Interestingly, however, in the alignment of viral and cellular members of the papain-like protease family, the only statistically significant relationship involving a viral protease was found to be between the RUB NS protease and several cellular proteases, possibly indicating that the RUB NS protease was only recently captured from cellular sources during virus evolution. However, from the same alignment, a resemblance was proposed between the alphavirus nsP2 protease and the potyvirus HC-Pro, a relationship which totally confuses origin-function similarities among these proteases. Resolution of the evolutionary relationship among the viral papain-like proteases will probably require definition of their three-dimensional structure.

Whatever the evolutionary origin of the RUB and alphavirus NS proteases, the apparent difference in their function has profound implications for the replication of the respective viruses. The processing of the alphavirus NSPs involves a sophisticated cascade mediated by the NS protease which regulates the course of virus RNA synthesis. In contrast, the apparent single *cis* cleavage mediated by the RUB NS protease would not allow this virus to evolve such an elaborate regulatory process.

ACKNOWLEDGMENTS

We thank Robert Simmons for technical imaging.

This work was supported by Public Health Service grants AI 21389 (T.K.F.) and AI 10793 (J.H.S.) from the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. **Baker, S. C., C. K. Shieh, M. F. Chang, D. M. Vannier, and M. M. C. Lai.** 1989. Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. J. Virol. **63:**3693–3699.
2. **Baker, S. C., K. Yokomori, S. Dong, R. Carlisle, A. E. Gorbalenya, E. V. Koonin, and M. M. C. Lai.** 1993. Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. J. Virol. **67:**6056–6063.
3. **Bonilla, P. J., S. A. Hughes, J. D. Pinon, and S. R. Weiss.** 1995. Characterization of the leader papain-like proteinase of MHV-A59: identification of a new in vitro cleavage site. Virology **209:**489–497.
4. **Bredenbeek, P. J., and C. M. Rice.** 1992. Animal RNA virus expression systems. Semin. Virol. **3:**297–310.

4a. **Carrington, J. C.** Personal communication.

5. **Carrington, J. C., S. M. Cary, T. D. Parks, and W. G. Dougherty.** 1989. A second protease encoded by a plant potyvirus. EMBO J. **8:**365–370.
6. **Carrington, J. C., and K. L. Herndon.** 1992. Characterization of the potyviral HC-Pro autoproteolytic cleavage site. Virology **187:**308–315.

6a. **Chen, J. P.** Unpublished data.

7. **Chen, J. P., and T. K. Frey.** 1995. Expression of the rubella virus structural proteins by an infectious Sindbis virus vector. Arch. Virol. **140:**2075–2084.
8. **Choi, G. H., D. M. Pawlyk, and D. L. Nuss.** 1991. The autocatalytic protease p29 encoded by a hypovirulence-associated virus of the chestnut blight fungus resembles the potyvirus-encoded protease HC-Pro. Virology **183:**747–752.
9. **de Groot, R. J., W. R. Hardy, Y. Shirako, and J. H. Strauss.** 1990. Cleavage-site preferences of Sindbis virus polyproteins containing the nonstructural proteinase: evidence for temporal regulation of polyprotein processing *in vivo*. EMBO J. **9:**2631–2638.
10. **de Groot, R. J., T. Rumenapf, R. J. Kuhn, E. G. Strauss, and J. H. Strauss.** 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. Proc. Natl. Acad. Sci. USA **88:**8967–8971.
11. **Dominguez, G., C.-Y. Wang, and T. K. Frey.** 1990. Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during Togavirus evolution. Virology **177:**225–238.
12. **Dong, S., and S. C. Baker.** 1994. Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. Virology **204:**541–549.
13. **Forng, R.-Y., and T. K. Frey.** 1995. Identification of the rubella virus nonstructural proteins. Virology **206:**843–853.
14. **Frey, T. K.** 1994. Molecular biology of rubella virus. Adv. Virus Res. **44:**69–160.
15. **Frey, T. K., and R. J. Jackson.** In vitro translation of the rubella virus nonstructural protein open reading frame. Submitted for publication.
16. **Gorbalenya, A. E., E. V. Koonin, and M. M.-C. Lai.** 1991. Putative papain-related thiol proteases of positive-strand RNA viruses. FEBS Lett. **288:**201–205.
17. **Hughes, S. A., P. J. Bonilla, and S. R. Weiss.** 1995. Identification of the murine coronavirus p28 cleavage site. J. Virol. **69:**809–813.
18. **Kim, J. C., R. A. Spence, P. F. Currier, X. Lu, and M. R. Denison.** 1995. Coronavirus protein processing and RNA synthesis is inhibited by the cysteine proteinase inhibitor E64d. Virology **208:**1–8.
19. **Kirchweger, R., E. Ziegler, B. J. Lamphear, D. Waters, H.-D. Liebig, W. Sommergruber, F. Sobrino, C. Hohenadl, D. Blaas, R. E. Rhoads, and T. Skern.** 1994. Foot-and-mouth disease virus leader proteinase: purification of the Lb form and determination of its cleavage site on eIF-4. J. Virol. **68:**5677–5684.
20. **Kleina, L. G., and M. J. Grubman.** 1992. Antiviral effects of a thiol protease inhibitor on foot-and-mouth disease virus. J. Virol. **66:**7168–7175.
21. **Koonin, E. V.** 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J. Gen. Virol. **72:**2197–2206.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
23. **Liljestrom, P., S. Lusa, D. Huylebroeck, and H. Garoff.** 1991. In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. J. Virol. **65:**4107–4113.
24. **Marr, L. D., C.-Y. Wang, and T. K. Frey.** 1994. Expression of the rubella virus nonstructural protein ORF and demonstration of proteolytic processing. Virology **198:**586–592.
25. **McGowan, E. B., E. Becker, and T. C. Detwiler.** 1989. Inhibition of papain in intact platelets by the thiol protease inhibitor E-64d. Biochem. Biophys. Res. Commun. **158:**432–435.
26. **Piccone, M. E., M. Zellner, T. F. Kumosinski, P. W. Mason, and M. J. Grubman.** 1995. Identification of the active-site residues of the L proteinase of foot-and-mouth disease virus. J. Virol. **69:**4950–4956.
27. **Pugachev, K. V., P. W. Mason, and T. K. Frey.** 1995. Sindbis vectors suppress secretion of subviral particles of Japanese encephalitis virus from mammalian cells infected with SIN-JEV recombinants. Virology **209:**155–166.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Shapira, R., and D. L. Nuss.** 1991. Gene expression by a hypovirulence-associated virus of the chestnut blight fungus involves 2 papain-like proteinase activities: essential residues and cleavage site requirements for P48 autoproteolysis. J. Biol. Chem. **266:**19419–19425.
30. **Shirako, Y., and J. H. Strauss.** 1990. Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. Virology **177:**54–64.
31. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1992. The 5′ end of the equine arteritis virus replicase gene encodes a papain-like cysteine protease. J. Virol. **66:**7040–7048.
32. **Strauss, E. G., R. J. de Groot, R. Levinson, and J. H. Strauss.** 1992. Identification of the active site residues in the nsP2 protease of Sindbis virus. Virology **191:**932–940.
33. **Strauss, J. H. (ed.).** 1990. Viral proteinases. Semin. Virol. **1:**307–384.
34. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, evolution. Microbiol. Rev. **58:**491–562.
35. **Wang, C.-Y., G. Dominguez, and T. K. Frey.** 1994. Construction of rubella virus genome-length cDNA clones and synthesis of infectious RNA transcripts. J. Virol. **68:**3550–3557.
36. **Weaver, S., A. Hagenbaugh, L. A. Bellew, S. Y. Netesov, V. E. Volchkov, G.-J. J. Chang, D. K. Clarke, L. Gousset, T. W. Scott, D. W. Trent, and J. J. Holland.** 1993. A comparison of the nucleotide sequences of Eastern and Western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses. Virology **197:**1–16.